Zurfer: Mobile Multimedia Access in Spatial, Social and Topical Context

Amy Hwang, Shane Ahern, Simon King, Mor Naaman, Rahul Nair, Jeannie Yang
Yahoo! Inc.
Berkeley, CA, USA
{amyhwang, sahern, simonk, mor, rnair, jeanniey}@yahoo-inc.com

ABSTRACT

What happens when you can access all the world's media, but the access is constrained by screen size, bandwidth, attention, and battery life? We present a novel mobile context-aware software prototype that enables access to images on the go. Our prototype utilizes the channel metaphor to give users contextual access to media of interest according to key dimensions: spatial, social, and topical.

Our experimental prototype attempts to be playful and simple to use, yet provide powerful and comprehensive media access. A temporally-driven sorting scheme for media items allows quick and easy access to items of interest in any dimension. For ad-hoc tasks, we extend the application with keyword search to deliver the long tail of media and images.

Elements of social interaction and communication around the photographs are built into the mobile application, to increase user engagement. The application utilizes Flickr.com as an image and social-network data source, but could easily be extended to support other websites and media formats.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

General Terms

human factors, design, algorithms

Keywords

photo browsing, mobile, context-aware, location-based

1. INTRODUCTION

Mobile devices are becoming more powerful and capable. "Smart Phones" are now programmable, network connected, support a screen resolution and colors that allow for rich media experiences, and include high speed processors. With these capabilities, the mobile phones in everybody's pockets

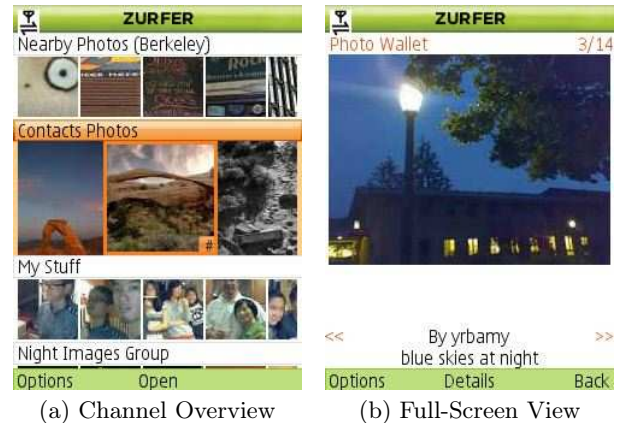

(a) Channel Overview   (b) Full-Screen View

Figure 1: Two screens from the Zurfer application: the channel overview screen, and the channel full-screen view.

are primed to become the future "photo wallet" – one that can enable access to not just a few select photos, but to all the world's media. However, some restrictions still apply for even the most advanced mobile phones on the market today. The devices' screens are small, and the interaction techniques are not as rich as those afforded by PCs and laptop computers. In addition, mobile settings often reduce the levels of attention and concentration on the devices interaction. Network bandwidth and battery life are some of the other issues that affect the functionality of mobile devices.

We present an experimental mobile software prototype, Zurfer[1], designed to support a mobile window to the world of multimedia. The design requirements for Zurfer included some potentially conflicting goals: playful and simple to use; enable common tasks; powerful methods to access the long tail of content. We based our design on a number of contextual dimensions that attempt to engage the users without requiring explicit input. In particular, we emphasize the spatial dimension as a way to elicit the users, via the mobile device, to engage and understand the physical world surrounding them.

The key ideas that are incorporated in the Zurfer prototype include:

- Generate and encourage engagement and serendipitous light-weight discovery, using a personalized and contex-

---

[1]http://zurfer.research.yahoo.com

tualized presentation of images. In our design, the main screen of the application includes the system's best guess at which images the users may be interested in, using two key dimensions: spatial (photos taken around the user) and social (photos from the users friends). To enhance the social interaction, we allow users to comment on photos, as well as read recent comments on their photos.

- Provide customization to allow for arbitrary interest and pin-pointed content. A separate configuration website allows users to select items of special interest (such as specific topics or social contacts). These selections appear in the main screen of the mobile application.

- Allow for key functions in respect to the user's own photo collection. Zurfer allows access to the user's entire photo collection via various dimensions and filters.

- Provide search mechanisms for unexpected interests and needs, and for "drilling down" from the default items.

The *channel* is the basic interaction metaphor in the Zurfer application. The main, opening screen of the application (Figure 1(a)) always displays several "basic channels" that are derived from the user's context - spatial, social and topical. An additional channel in the main screen provides easy access to the user's own photos. Finally, the users can add "custom channels" to the main screen according to their own personal interests and needs. Our mobile application is tied to the popular image-sharing website, Flickr. Zurfer uses Flickr as a backend for images, social network data, user information and registration and so forth. The prototype is not affiliated with the Flickr website, making the integration incomplete. However, we believe that even for an experimental research prototype, leveraging the active community on Flickr instead of building our own online image storing and sharing service is advantageous for many reasons.

Zurfer is designed to run on current smart phones, with no special interaction requirement such as stylus.

2. RELATED WORK

In this section we provide a short overview of previous mobile and context-aware media browsers. We do not venture into related user-research and ethnography work, and also limit the discussion here to image content (not video).

Like our work, the prototype of Pauty et al. [8] is both mobile and location-aware. The authors present a mostly location-driven browser that displays photos taken around the user in various directions and allows the users to virtually explore their surroundings. The system does not offer a way to handle clutter caused by the exceedingly large number of photos available online. In this work, we use a novel method to extract important concepts that are nearby and show them to the users; in addition, we place an emphasis on recent "nearby" activity, pushing a social-spatial-temporal angle that promises to be more engaging.

Enabling image browsing on small-screen devices was the topic of several research efforts [2, 3, 6, 9]. Most of these efforts assume that the mobile images are stored locally on the mobile device, or the existence of a fixed image collection that the user can browse (e.g., the user's personal image collection). The focus of [2, 3] is the interaction and visualization techniques, while [6, 9] focus on automatic analysis of the data to a form that is useful for mobile browsing (but do not investigate the details of the interaction). Our focus here is experimentally extending the available set of photos beyond the personal realm, and allowing access to arbitrary photos in key dimensions (spatial and social).

To better handle the small screen size, a number of projects (e.g., [4, 10]) considered automatically detecting the "salient" features and areas in an image and using that knowledge when displaying the image on small devices. These techniques are orthogonal to our work and can be used to improve the display and browsing experience.

Location-based browsing of images has been a topic of interest since 2003. Toyama et al. [11] used a PC-based map-and-timeline interface to allow for browsing of geo-referenced images. Naaman et al. [5] had used different context cues to assist in image browsing; O'hara et al. [7] later used similar ideas in their work. None of these applications had considered small screen access to location-based images.

3. ZURFER INTERACTION BASICS

The basic interaction metaphor in our experimental interface is the "channel". Figure 1(a) shows the main screen for Zurfer. In this view, each row of images represents a single channel. The user can navigate up and down the rows to view the different image channels. When a channel is highlighted (like the 'Contacts' channel in the figure), the user can browse images in the channel by navigating forward or backwards (right and left). Most channels are sorted by time, and the most recent photos are shown first. In theory, each channel can have an unbounded number of images – more images get loaded as the user browses through the list. The 4-way interaction (up, down, right, left) is simply done using the 4-way navigation button available on most phones today.

An important design consideration when using the channel metaphor, was to create a feel of "Infinite scrolling" in the Zurfer channels (as opposed to "paging" through a set of results for each channel). In other words, as long as there are available images in the channel, the user can keep navigating forward without any significant delay; the images are loaded in the background as needed.

Selecting a channel (by clicking into the 4-way navigation button) brings up more options within that channel. These options often include a further breakdown of the channel into "sub-channels" and an option to search the content of that channel. The presentation of sub-channels is equivalent to the main channel screen in Figure 1(a).

The default channel view, as shown in Figure 1(a), is optimized for showing several images on screen, as long as the thumbnails are big enough to be of some use (the thumbnail sizes depend on the screen resolution). At any point, the user can enlarge the images for an almost-full-screen view, as shown in Figure 1(b). In the full-screen view, users can still navigate forward and backwards in the channel.

The users are given the option of commenting on any photo they see as well as the ability to mark as favorite any photos that are not their own. In the details page, accessible from the full-screen view of each photo, the users can also view the current comments for the photo and additional photo metadata such as time, tags, descriptions and so forth.

4. BASIC CHANNELS

In this section we describe the main channels that are shown by default to every user of the Zurfer mobile proto-

type. These basic channels are designed to supply as much functionality and engagement via discovery as possible. The channels are driven by, but not limited to, the spatial and social context of the user.

4.1 Nearby photos Channel

The first displayed channel is the `Nearby Photos` channel. Zurfer detects the user's location using the cell tower ID to which the phone is currently connected; alternatively, the user can enter their location manually. The location resolution, therefore, ranges from exact address to city level.

For each location, Zurfer gets the latest photos from Flickr that were marked as taken at that location. The goal of this selection is not virtual navigation of the physical environment, but rather serendipitous discovery of current activity around the user.

The `Nearby Photos` channel can also be expanded to show "local highlights". These are key landmarks around the user's location that could serve as visit or photo recommendations and are automatically extracted from the Flickr data for each location (as described in [1]). For example, for a user in zipcode 10025 (northwest of Central Park, Manhattan, New York City) the local highlights chosen and shown by Zurfer include Harlem, Guggenheim, Central Park, Reservoir, St. John the Divine and Broadway. The presentation of the local highlights is in channel format, just like Figure 1(a), with each highlight as a separate channel.

Also in the breakdown of the `Nearby Photos` channel, Zurfer allows the user quick access to their own photos from that location, perhaps including photos from past events that the user attended.

Key experimental questions for this channel include whether users engage with location-based photographs from other users; whether the local highlights view is useful, whether access to the user's own photos from a location is effective and under which conditions, etc.

4.2 Social Channels

The social channels include the `Recent Comments` channel and the `Contacts' Photos` channel.

`Recent Comments` is a socially-driven channel that is geared towards engagement and communication. This channel includes all the photos by the user that were commented on in the last 24 hours. By allowing the user to view the recent comments while mobile, we hope to create a social stickiness, encouraging the user to use Zurfer often to check on new activity. The `Recent Comments` channel also includes activity on photos from other Flickr photographers that the user commented on. These photos are shown to promote discussion – a user would be able to quickly view comments added in response to their own, and respond if needed.

The `Contacts' Photos` channel shows the most recent photos upload by the user's contacts (their social network connections on Flickr). This channel represents the key social dimension, giving the user a quick overview of their friend's activities and sometimes even whereabouts. The `Contacts' Photos` channel can be expanded into a screen of individual contact channels, where each channel includes images from a single contact.

Key questions for the social channels in our experimental prototype include when, and how users are engaged with social contact's photos; who enters or reads comments while mobile, and so forth.

4.3 My Stuff Channel

The `My Stuff` channel includes all the user's photos, as well as special functionality to enable quick access to the user's most important images. The default set of images in the `My Stuff` channel is all the photos in the user's Flickr collection, ordered temporally such that the latest photos are shown first. The channel expansion, though, includes a number of features. The first is a breakdown of the user's photos by the recent locations where the user was taking photos, and a breakdown by the important recent tags the user has used for their images.

More interestingly, one of the sub-channels under `My Stuff` is the `Photo Wallet` channel. The physical photo wallet usually includes special photos the user carries around with them at all time (say, photos of the kids) for co-present sharing or individual reflection. Similarly, in Zurfer's `Photo Wallet` the user can save any of their photos for quick access later. Whenever the user views one of their photos on Zurfer, they have the option to add that photo to the Wallet (they can also do so on the web).

Finally, `My Stuff` includes a sub-channel for the user's "favorite" photos – photos taken by user Flickr members that the user had marked as "favorites". The idea is to enable the user to get quick access on Zurfer to any photo they liked on Flickr, again to be used for sharing or reflection.

Key experimental questions for the `My Stuff` channel include what photos of the personal collection are accessed and under which circumstances.

4.4 Topical Channels

The basic channels include one generic "topical" channel, `Interesting Today`. This channel includes photos that may be of topical interest but are not necessarily derived from the user's context. Of course, we could attempt to guess the user's topical interests based on their Flickr activity (e.g., tags they use for their photos). However, we opt in Zurfer for one generic topical channel, and allow customization for the user to state their own interests (see Section 5).

The `Interesting Today` channel includes photos that rank high in Flickr "interestingness" (a metric that roughly corresponds to photos with many views and comments). It also includes photos of a "topic of the day" which is selected daily by our research team. All in all, the `Interesting Today` channel is geared to provide the user with an interesting and visually-compelling image discovery experience.

In summary, the basic channels that are included in Zurfer's main channel view allow easy access (via simple 4-way navigation) in a number of key dimensions: spatial (recent activity and landmarks around the user), social (activity around the user's photos and the latest photos from their contacts), and topical (interesting photos and random visually compelling photos). In addition, one of the basic channels provides complete access to the user's own personal collection on Flickr. If all this is not enough, a user can add other channels of interest to their main channel page. We discuss this option next.

5. CUSTOM CHANNELS

Zurfer allows the user to choose what additional channels they would like to see in their mobile photo application. These additional channels are shown as Main channels in the main application screen.
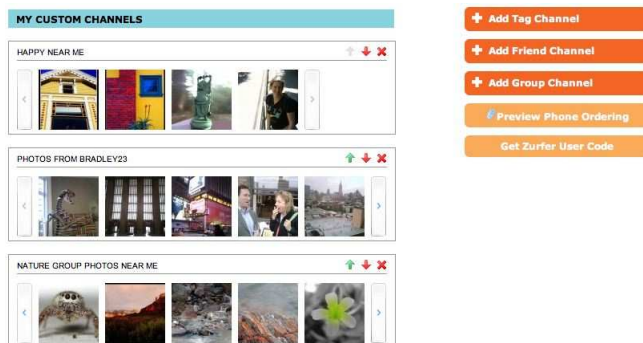
Figure 2: A partial screenshot of Zurfer's channel configuration page, including three user-configured channels.

Users can add a "custom channel" that correspond to a Flickr tag, a Flickr user, or a Flickr group.[2] For example, the "puppy" tag channel will always show the latest Flickr photos tagged "puppy". A Flickr user channel will show the latest photos from that user, and a group channel will show the latest from that group.

Each of these channels can be spatially-filtered, to show only photos from the user's current location. For example, a user interested in graffiti can create a channel that shows all the graffiti photos taken around them. Part of the configuration screen and a few sample user-selected channels are shown in Figure 2.

In our experimental prototype, we are interested in exploring the types of channels that are added by users to their Zurfer application, and how these are used.

6. OPTIMIZATION

A number of optimization measures were implemented in Zurfer to create the best possible mobile experience, including pre-fetching, caching, automatic thumbnail generation, and server-side channel logic. Pre-fetching is implemented such that the user does not have to wait for images to download. After fetching images that are in current view, Zurfer automatically fetches images that the user is likely to navigate to – e.g., photos in the forward direction of the currently-viewed channel. Additionally, Zurfer often loads alternate images in the channel (every other photo) – opting for a better sampling of channel content when users browse forwards quickly.

Caching is optional in Zurfer, and simply involved saving a local copy of viewed images on the phone's memory. Caching can reduce the network load as well as the image wait time. However, as network speeds improve, we do not see great benefits to caching images on the phone.

Finally, the Zurfer server is responsible for generating the thumbnails so that the costly thumbnail generation computation is not done on the phone. Similarly, all the logic of channel generation is on the server.

7. USER STUDY

A short user study was performed in an early stage of the prototype. The main conclusion of the study, performed

---

[2]A Flickr group is a pool of photos (usually with a common subject) to which multiple users can contribute.

in lab settings with five users, were that user interests in regards to our experimental prototype varied. Some users mainly expressed interest in the social aspect of the prototype. Others were more interested in the spatially-driven information. When asked about location-based interests, most users expressed the need to select their own topics of interest for nearby photos. These preferences may, change in a real deployment outside the lab settings.

Other than the dimensions of interest, our preliminary study demonstrated that the channel metaphor was clear and intuitive for the users. The basic details of the interaction were also straightforward, as well as the meaning of the basic channels.

8. CONCLUSIONS

We have described Zurfer, a novel, experimental mobile photo browser that is designed to engage the user in the spatial, social and topical browsing dimensions. The user interaction is based on a channel metaphor that allows users to intuitively and simultaneously view multiple streams of information. We intend to perform a wider deployment that will allow us to investigate the salient dimensions of browsing images in mobile settings, and verify Zurfer's experimental features.

9. REFERENCES

[1] S. Ahern, M. Naaman, R. Nair, and J. Yang. World explorer: Visualizing aggregate data from unstructured text in geo-referenced collections. In *Proc. JCDL 2007*.

[2] S. Harada, M. Naaman, Y. J. Song, Q. Wang, and A. Paepcke. Lost in memories: Interacting with photo collections on PDAs. In *Proc. JCDL 2004*.

[3] A. Khella and B. B. Bederson. Pocket PhotoMesa: a zoomable image browser for PDAs. In *Proc. MUM 2004*.

[4] H. Liu, X. Xie, W.-Y. Ma, and H.-J. Zhang. Automatic browsing of large pictures on mobile devices. In *Proc. ACM Multimedia 2003*.

[5] M. Naaman, S. Harada, Q. Wang, H. Garcia-Molina, and A. Paepcke. Context data in geo-referenced digital photo collections. In *Proc. ACM Multimedia 2004*

[6] M. Naaman, Y. J. Song, A. Paepcke, and H. Garcia-Molina. Automatic organization for digital photographs with geographic coordinates. In *Proc. JCDL 2004*.

[7] N. O'Hare, C. Gurrin, H. Lee, N. Murphy, A. F. Smeaton, and G. J. Jones. My digital photos: Where and when? In *Proc. ACM Multimedia 2005*.

[8] J. Pauty, P. Couderc, and M. Banâtre. Using context to navigate through a photo collection. In *Proc. MobileHCI 2005*.

[9] A. Pigeau and M. Gelgon. Building and tracking hierarchical geographical & temporal partitions for image collection management on mobile devices. In *Proc. ACM Multimedia 2005*.

[10] V. Setlur, S. Takagi, R. Raskar, M. Gleicher, and B. Gooch. Automatic image retargeting. In *Proc. MUM 2005*.

[11] K. Toyama, R. Logan, and A. Roseway. Geographic location tags on digital images. In *Proc. ACM Multimedia 2003*.